

McNamara Declaration

Exhibit 46

part 1

Internet Archive Open Libraries Proposal

MacArthur Foundation 100&Change



Table of Contents

| | |
|---------------------------------------------------------|-----------|
| A. OVERVIEW | 5 |
| Executive Summary | 6 |
| Location for Proposed Project | 8 |
| Detailed Project Plan | 9 |
| B. COMMUNITY ENGAGEMENT | 31 |
| Stakeholder Identification | 32 |
| Inclusiveness | 34 |
| Persons With Disabilities | 35 |
| Stakeholder Engagement | 36 |
| Stakeholder Influence | 38 |
| C. SCALABILITY OF THE PROJECT | 40 |
| Planning for Scale | 41 |
| Scaling Technologies | 41 |
| Expanding Diversity | 42 |
| The Credibility of the Proposed Solution | 43 |
| Support for the Change | 44 |
| Advantage of the Proposed Solution | 45 |
| Ease of Transferring and Applying the Solution at Scale | 47 |
| Organizational Capacity to Implement at Scale | 49 |
| Financial Sustainability of the Proposed Solution | 50 |

| | |
|------------------------------------------------|----|
| D. MONITORING, EVALUATION, AND LEARNING | 52 |
| MONITORING | 53 |
| Results | 53 |
| Tracking | 55 |
| Milestones | 57 |
| EVALUATIONS | 60 |
| Changes for Beneficiaries | 60 |
| Data Sources | 62 |
| Methodology | 63 |
| Gender Analysis | 65 |
| LEARNING | 66 |
| Findings | 66 |
| Dissemination | 68 |
| E. ADDITIONAL INFORMATION | 70 |
| Detailed Budget Narrative | 71 |
| Project Risk Management | 73 |
| F. FINANCIAL STRENGTH AND STABILITY | 76 |
| G. TECHNICAL | 81 |
| Technical Explanation | 82 |
| Conclusion | 86 |



Open Libraries

a project by the Internet Archive

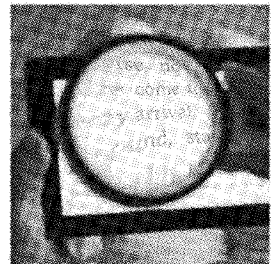
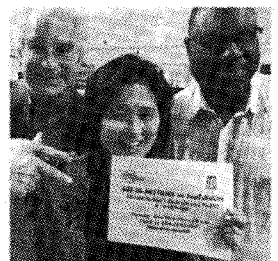
Internet Archive

100&Change

| | |
|-------------------------------|-----------|
| H. ANCILLARY MATERIALS | 87 |
| Evidence of Engagement | 88 |
| Biographies of Key Staff | 187 |
| Bibliography | 226 |

* Website: <http://openlibraries.online>

A. Overview



Executive Summary

Looking for a trusted source of information? For millions of citizens there's only one place to head: their local library—society's great equalizer, where Wi-Fi, computers, and knowledge are offered for free, to anyone entering the door. Yet, due to distance, time, cost or disability, people in marginalized populations are too often denied access to physical books. Today for many learners, if a book isn't digital, it's as if it doesn't exist. Yet there's almost a century of knowledge still living only on the printed page, missing from our digital shelves.

The Internet Archive's *Open Libraries* offers a solution bringing four million books online, through purchase or digitization, while honoring the rights of creators and expanding their online reach. Added to our existing 2.5 million ebooks, we can build the online equivalent of a great, modern public library. Working with US libraries and Benetech, operator of the world's largest digital library for people with disabilities that impact reading,¹ the Internet Archive (IA) will bring millions of free digital books to billions of people. For the blind, ebooks are a lifeline, yet less than one in ten exists in accessible formats.² Digital content becomes instantly available to people in rural areas, with widely ranging physical abilities. By digitizing millions of books, we unlock them for communities with limited or no access.

At the same time, we have a rare opportunity to shape digital collections serving communities that are increasingly diverse. In 2015, 50% of newborn US babies were children of color,³ yet over a twenty-year span, on average only 10% of children's books contained multicultural content.⁴ Because our library shelves can and must be as diverse as our readers, we will curate inclusive content.

IA will select and preserve diverse collections and help libraries greatly expand their digital holdings. In 2013, ebooks comprised an average of 17% of US public libraries collections;⁵ we can turn 80% of library collections digital by 2023. We'll build a financially sustainable infrastructure for at-scale ebook circulation so US libraries that own the hardcopy can offer their patrons temporary digital access, just like loaning a book. Working with Benetech, we'll expand ebooks for people with disabilities by 10x and share these books across 29 nations through global copyright agreements.⁶

In this era of disinformation, ready access to trustworthy sources is critical. Library books are trusted sources for lifelong learning. By bringing them online, we empower journalists, educators and Wikipedia editors to cite "snippets" directly, grounding readers in the vetted, published record.

A century ago, Andrew Carnegie funded a vast network of public libraries because he recognized democracy can only exist when citizens have equal access to diverse information. Libraries continue to play that vital role, welcoming the whole of society to use their free resources for individual learning, while respecting readers' privacy and dignity. Through its support, the MacArthur Foundation can help us build an enduring asset for libraries across this nation, ensuring that all citizens—including our most vulnerable—have equal and unfettered access to knowledge.

Location for Proposed Project

We will not be conducting any work in nations that have been sanctioned by the US government.

Internet Archive staff members located in the United States will do the primary technology development for *Open Libraries*. We will digitize and re-publish approximately 3-4 million books at Internet Archive-run facilities in the United States, Hong Kong, Shenzhen, China, and possibly at a new center in Asia or Latin America. Book curation teams will be based in the United States with input from accessibility partners in the (currently) 29 nations that have ratified the “Marrakesh Treaty to Facilitate Access to Published Works by Visually Impaired Persons and Persons with Print Disabilities.”¹ Our current dissemination plan for the print disabled begins with India, Canada, and the 12 Latin American nations that have ratified the Marrakesh Treaty: Argentina, Brazil, Chile, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Panama, Paraguay, Peru, Uruguay.² (However, all of our texts will be available to people with disabilities worldwide.)

Through Benetech, we will remediate 10,000+ books to higher accessibility by working with their contractors in India, Laos, and Kenya.

Five-Year Plan for *Open Libraries*— a project by the Internet Archive

At the **Internet Archive**, we believe passionately that access to knowledge is a fundamental human right. Knowledge makes us stronger and more resilient; it provides people pathways to education and the means to secure a job. But many people in marginalized communities encounter daunting barriers to knowledge. The Accessible Books Consortium states that for the blind, “the lack of accessible books is a very real barrier to getting an education and leading an independent, productive life.”¹ By harnessing scalable technologies, social entrepreneurial capital, and leaders in the field who support this change, *Open Libraries* seeks to create a lifeline to trusted information by bringing millions of digital books to billions of learners around the world.

How will we do this? Through one of society’s most respected and trusted institutions—the library. In the United States, libraries have long ceased to be places that just lend books. If you haven’t been to a public library lately, you might be amazed. Large crowds stand outside each morning, waiting for the doors to open. Inside there are makerspaces, media labs, and massive open online courses (MOOCs). In U.S. cities, libraries have become front line providers of community services, from free Wi-Fi to literacy campaigns to citizenship classes for immigrants. Surveys show that low-income Americans, Latinos, and African Americans are the most likely to say closing libraries would have a major impact on their lives and communities.² By strengthening libraries we serve the needs of America’s most vulnerable populations.

1 Accessible Books Consortium Bringing books to persons with print disabilities
<https://publishsa.co.za/file/1446643888wec-abc-brochure.pdf>

2 Libraries at the Crossroad Pew Research Center September 2015 <http://www.pewinternet.org/2015/09/15/libraries-at-the-crossroads/>



Internet Archive

100&Change

Five-Year Plan for *Open Libraries*— a project by the Internet Archive

The Internet Archive will help to transform the US library system by providing digital delivery of 80% of the average library's collection by 2023, reducing the barriers to information created by distance, time-constraints, cost, and disability. By leveraging this one-time investment in digitization, libraries may save as much as \$3 billion in interlibrary loan fees over the next decade. Ebooks also free up shelf space, allowing librarians to buy new books from publishers and divert their valuable real estate to serve their core mission: to enrich, educate, and empower every individual in our diverse communities.³

Internet Archive's *Open Libraries* project will:

- **CURATE** diverse & inclusive collections
- **INCREASE ACCESSIBLE CONTENT** by 10x for all people with print disabilities
- **SCAN & PROVIDE DIGITAL ACCESS** to books through libraries that own the hard copy
- **ENSURE READER PRIVACY & LONG TERM PUBLIC ACCESS**
- **PRESERVE** millions of books now missing from our digital shelves

Here is our Project Plan: (Please see attached Worksheets)

A. Governance

The Internet Archive has rigorous fiscal policies and oversight to ensure our projects meet deliverables on time and within budget. We have a track record of successfully managing projects of enormous scale and marshalling our resources efficiently. *Open Libraries* will benefit from the guidance and oversight of several governing boards. As reflected in **Section A**, the Internet Archive Board of Directors (#2), our outside auditors (#3), and *Open Libraries*

³ From the Mission Statement of the Los Angeles Public Library: <https://www.lapl.org/about-lapl/about-library>

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

25-member Advisory Board (#1) will each meet annually to review the project's progress, and ensure our financial and operational accountability. Our Advisors will form seven working groups in Curation, Accessibility, Technology, Collection Development, Service Plan Development, Legal, and Financial Modeling. These global leaders in finance, law, libraries, and education will greatly inform the way we conceptualize and execute our plans.

Internally, our Finance Team and grant manager will conduct quarterly reviews against our benchmarks (#4-5). As a 501(c)(3), the Internet Archive has successfully completed IRS-required audits since 2006. Governance will be overseen by **Digital Librarian, Brewster Kahle**.

B. Building the *Open Libraries* Team

Internet Archive's staff is comprised of 164 employees and 18 contract employees, and we are in the process of hiring 12 additional staff. For this project we will add 25 new positions over the five-year grant period, a 14% increase. To prepare for this rapid round of hiring, our **Director of Human Resources, BZ Petroff** has spent the last six months implementing new software solutions to streamline our human capital management processes, from recruiting through onboarding. (Petroff has a stellar track record as an HR leader, once overseeing 35 recruitment consultants hiring 700 new employees for Lucasfilm in a single year). She will bring on an outside recruitment firm and one staff recruitment specialist to help us find, attract, and hire talented, mission-driven team members. Our brand: people who are passionate—about their families, their interests, their communities—and bring that passion to their work.

C. Curating Books for Maximum Impact

One of this project's most exciting opportunities is to select which four million books—the equivalent of a major urban public library—to carry

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

forward for today's online learners as well as future generations. We had to ask ourselves: How do you curate millions of books in a way that is inclusive, responsive to the community and transparent? **Section C** outlines our approach. We start with core lists serving diverse segments of our society: K-16 students, public library patrons, people with disabilities, Spanish-speakers. Curation at this scale is mostly about parsing data. Take, for example, how we might derive the College Core list. Our Curator of Books will start by taking the metadata (ISBNs) from a typical mid-size college library collection, and cross it with the Open Syllabus list of the books most-assigned in college classes. We could then run "scripts" to see which of these books are already in IA's collections. Our curator might analyze the books that remain for overlap with those most widely held by the 16,000 libraries who subscribe to OCLC's cataloging service. Doing these overlap studies gives us a prioritized wish list of books highly relevant to college students. But computers alone cannot build a good collection. That's why we are working with the Digital Public Library of America (DPLA), whose Curation Corps of 20 volunteer librarians will help hand-select a total of 100,000 books in Year One, and one million over the life of the project. These experienced librarians will help select titles reflecting our diverse communities (#6).

All of these wish lists will be posted on the new dashboard built by the Digital Library Federation (DLF), along with best practices in selecting culturally diverse books. DLF will lead a transparent process for community feedback. These activities are listed in #8, as the Inclusive Curation Project and overseen by **IA's Curator of Books**.

D. Book Sourcing—Developing a Pipeline of Books

Selecting the right books is only half the challenge. Next we have to find them. That's the job of the **Director of Books**: procuring the books on our

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

curated lists. We will buy as many ebooks as we can from publishers and authors (#2). Others we will buy from booksellers as hard copies and digitize them. In many cases, libraries will give us an extra copy of the books that they are weeding. Our plan assumes that 585,000 books will come from partners who will pay to have their material digitized (#1). The third category is publishers such as MIT Press who are willing to digitize and lend their backlists (#3), in order to reach more people. Arcadia has agreed to underwrite this effort to digitize the historic output of university presses with a \$1 million grant.

We will also support the many global nonprofits who produce copyright-free books, inviting them to join *Open Libraries* to add their culturally-appropriate, high value texts to our collections (#4). We can ensure that their legacy of rich content is not lost to history or left to the exigencies of commercial services. Our friends at Worldreader got their start in our office space, with just a handful of staff. We hope *Open Libraries* can support the work of all social good organizations providing access to knowledge around the world. Our **Books Business Development Manager** will oversee this scope of work.

E. Technology Plan

Here we broadly outline the technologies we will build or license for *Open Libraries*. Technology development is rarely a straight line, so like any successful plan, we will remain iterative and flexible.

1. Integrate ebooks purchased from or donated by publishers and authors

Wherever possible, we will purchase books that have been authored and published in a digital format (ebooks). When policy and agreements allow, these “born digital” works will be added to the overall collection with minimal technical processing.

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

2. *Increase book digitization capacity*

The Internet Archive currently digitizes approximately 1,000 books per workday (250,000 per year). This project requires expanding our digitization capacity to peak at 750,000 books in 2020, 2021, and 2022. We will be refining our own hardware, the TableTop Scribe scanner, designed and built by the Internet Archive for this express purpose. Ordering parts, assembling, and testing the hardware requires planning and lead time. Additionally, we may need to refine and expand the capacity of the software infrastructure that helps convert the page images into digital books.

3. *Improve the accessibility of Internet Archive books*

Shifting books into digital formats is the first step to making them accessible for the approximately 10% of all people who have disabilities that impact reading. Once digital, book files can be magnified, automatically read aloud (using synthesized speech), or used with special software, according to the needs of each individual.

As part of our engagement with this community, we have learned that our technology and methods for producing accessible texts can be improved. Consequently, we propose to:

- improve our digitization processes for new materials
- selectively enhance some digitized materials to bring them to a higher level of accessibility

We are already engaging with Benetech (operator of the world's largest accessible library, Bookshare), National Federation of the Blind (NFB), LightHouse for the Blind, US Library of Congress' National Library Service (NLS) for the Blind and Physically Handicapped, Accessible Digital Library of India, Centre for Internet and Society (India), and Learning Ally. For many years we've partnered with the accessibility teams at universities, including the Scholars Portal for Ontario (led by the Accessibility Librarian for

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

University of Toronto Libraries), and the Accessible Media Services group at the University of Illinois.

4. *Improve the accessibility of our websites and bookreader*

Our plans include assessing the accessibility of our websites, archive.org and OpenLibrary.org. Where necessary, we will improve the websites to support better engagement, working closely with our partners who have deep experience serving in the practical needs of learners with disabilities.

5. *Increase circulation of ebooks to global communities of people with disabilities*

This project makes books accessible and available to individuals with disabilities under current laws and treaties. Delivering to print-disabled populations requires both good-quality digital versions of materials (ebooks) and establishing the partnerships and distribution systems essential to getting these ebooks into their hands and reading devices.

There are three significant areas and phases for this expansion:

a. *Work with US providers of services to the blind to authorize large-scale access*

In the United States, there are several organizations chartered to provide accessible reading materials to blind and print-disabled communities, each with a significant user base and distinct online portal. The Chafee Amendment (Section 121 of the Copyright Act⁴) enables us to offer full download of any in-copyright text to any person with a certified print disability. We are working with the top four US organizations to authorize their members for full access to our millions of books.

We are partnering with teams from Benetech, the Library of Congress National Library Service (NLS) who operates Braille and Audio Reading Download (BARD), Learning Ally, and the National Federation of the Blind.

⁴ 17 U.S. Code § 121 - Limitations on exclusive rights: Reproduction for blind or other people with disabilities
<https://www.law.cornell.edu/uscode/text/17/121>

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

A first step requires integration with each organization's user authorization system, allowing IA to recognize any print-disabled user member and authorize them to access any of our digitized texts in a format appropriate for adaptive devices. These organizations have already provided technical information and offered to help connect their user bases to archive.org. If you combined the top three accessible libraries in the United States, the size of the combined collection would be approximately 699,000 titles (538,000 at Benetech, 81,000 at NLS, and 80,000 at Learning Ally) and that aggregation certainly includes significant overlap. We propose bringing more than 6.5 million titles to this community of users, offering them the breadth and depth of topics, authors, and experiences enjoyed by those without the same disabilities. Our plan also includes working with Benetech to bring a special collection of 10-15,000 highly requested books to even fuller accessibility.

b. Expansion to selected Marrakesh Treaty ratifying nations

With digital books, we can serve blind communities far beyond the US borders. The Marrakesh Treaty⁵ enables ratifying nations to share published materials with each other. The provisions are currently being considered by the US Congress, however Canada has already ratified this global intellectual property treaty. The Internet Archive Canada will engage with targeted Marrakesh signatories to make digitized books available to their print-disabled communities. We will begin in India, home to the world's largest blind population.

c. International expansion to Marrakesh nations on a regional basis, starting with Spanish-speaking countries of Central and South America.

We will support our partner, Benetech, to create and publicize a regional, language-centric portal for integration with organizations serving the blind

⁵ World Intellectual Property Organizations, Marrakesh Treaty http://www.wipo.int/treaties/en/ShowResults.jsp?lang=en&treaty_id=843

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

of any Spanish-speaking Marrakesh nation. This portal will serve Spanish-speaking nations in Central and South-America (12 of which have ratified the Marrakesh Treaty) and ensure the bidirectional sharing of all digital books, including and especially those available in Spanish. Ten of the 12 nations do not have any national body serving the needs of people with print disabilities. This portal will fill an enormous gap.

6. *Expand circulation and lending capacity*

Through OpenLibrary.org we have six years of experience lending digital books on behalf of 100+ partner libraries. This project proposes to expand significantly the infrastructure and services required to bring the collections of thousands of libraries online for lending, using tools that can be managed by the libraries themselves.

Our partners have made it clear: they need a turnkey solution that requires minimal technical investment on their part, so IA is developing a web-based portal that allows for easy collection curation by library administrators. Further, we will work to help integrate digital books with libraries' existing discovery systems, such as online card catalogs, subject collections, and lists of staff favorites.

To help augment libraries' existing digital lending programs, we will also build circulation systems for careful control over digital access. We will staff a partnership support team providing library staff with the necessary training for initial setup and ongoing curation.

7. *Enable hyperlinks in citations to resolve to snippets of text*

In a time of growing disinformation, knowing your sources really matters. Ideally all books cited in Wikipedia, educator's lesson plans, and in news and journal articles would become hypertext links to those books, allowing the reader to click and open to the right passage or "snippet" for easy consumption. If readers want to read more, then they may need to borrow the book, a process we will attempt to simplify. By weaving large swaths

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

of published works into the web itself, we will be fulfilling one of the great promises of the hypertext web, to link concepts directly to their sources.

We will work toward this goal by creating stable URLs for pages into books, build translation tools for standard footnote formats to those URLs, and then work with Wikipedia, OER Commons, and others to integrate these new capabilities into editors' workflows.

Since 2013, we have worked with Wikipedia communities to automatically “heal” more than two million broken links in their articles, replacing them with links to the original page that we’ve archived in the Wayback Machine. Working with the 286 disparate communities⁶ that make up Wikipedia has taken time, but we feel confident we can achieve integration with *Open Libraries*' books as well.

8. *Enable research use of the digital book dataset*

A digital repository of more than four million books and related metadata creates a data set of great interest to researchers and data scientists. Researchers may derive significant scholarly benefit from analyzing data at this scale—such as the analysis of linguistic patterns or AI-enhanced analysis of historical trends.

We will enable what our profession calls “non-consumptive” research use—research that does not include reading a book or displaying large portions of in-copyright texts, but performs computational analysis of book data. Realizing this outcome will require us to build secure procedures and appropriate hosting that supports “bulk access” to book data at scale.

F. Building Partnerships with Key Stakeholders

We will achieve scale and impact by building partnerships throughout the

⁶ Wikimedia Foundation, List of Wikipedias https://meta.wikimedia.org/wiki/List_of_Wikipedias

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

entire ecosystem of libraries. That requires validation and support from key stakeholders. It means engaging in a community-centric process to design tools and services. This process is being led by **Director of Partnerships, Wendy Hanamura** and **Sr. Strategist, Jim Michalko**. As outlined in **Section F**, we've secured early endorsements from key stakeholders (#6). The boards of the largest associations representing libraries—the American Library Association (ALA), Public Library Association (PLA), and Chief Officers of State Library Agencies (COSLA)—support this effort. So do many authors and publishers. We are working with the Authors Alliance and MIT Press to align around our common goal: creating a new generation of readers and critical thinkers who love books.

Our scaling strategy involves engaging innovators from a broad range of archetypes to serve as first movers, championing *Open Libraries* within their associations and networks (#5). Our first calls were to educators and knowledge providers (#3) in the movement to adopt Open Educational Resources (OER). They will connect our resources to the learners who need them most—students who cannot afford textbooks, veterans getting their GEDs, and independent scholars who don't have the resources of a major university.

Our partners are building this service side-by-side with us. These include the curation specialists at the Digital Public Library of America (DPLA), the Digital Library Federation (DLF), and Benetech, operator of the world's largest accessible library, serving people in 70 countries (#1). We will see network effects when we integrate with the technologies of key service providers (#4) that already support thousands of libraries around the world. Likewise, we can reach millions of people with print disabilities by integrating

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

services with the leading organizations in that sector (#2).

G. Establishing Legal Frameworks and Agreements

The goal of *Open Libraries* is to forge a path that honors the rights of creators and expands their online reach, while serving the compelling interests of society, especially our most marginalized communities. As Georgetown's Associate Dean, Michelle Wu, explained in her presentation to MacArthur Foundation staff, "I'm here to address the lack of understanding about copyright law that I think has held libraries back for decades. Many do not realize that Congress enacted equally strong protections for society by way of fair use, which recognizes that under certain conditions, actions that would be copyright infringement are permitted as equitable and in the public interest." Clarifying this legal framework for the entire ecosystem of university counsel, public library leaders, authors, and publishers would be a major contribution to society. Without it, we are stuck in false opposition, afraid to challenge the status quo, poorly serving the interests of learners of every hue.

We have now gathered a cadre of library leaders willing to share the risk of legal opposition in order to move society forward. Refining a common understanding of fair use and digital library lending is a fundamental step to legitimizing this transformational system-wide change.

In May, our advisor, Pam Samuelson, one of the nation's preeminent copyright scholars, came to us with a proposition. She offered to convene a day-long meeting with 20 of America's most respected fair use experts. They would consider our plans, offer guidance about fair use in general, and form a working group willing to craft a joint statement for others to consider. This work is now in progress with plans to continue to gather the working group annually (#2) and disseminate their joint statement through events such as

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

“Fair Use Town Halls,” proposed by Harvard University copyright advisor, Kyle Courtney, as part of his global Fair Use Week activities in 2018 (#1). This critical area of work is led by **Internet Archive’s legal counsel, Lila Bailey**, who has spent her career defending the reader’s rights to privacy and specializing in digital copyright law.

H. Establish Systems to Inventory, Track, and Store our Physical Assets

For the Internet Archive, reducing costs means controlling the supply chain at every point of the process. That’s why we invented our own scanners, run our own data centers, and even staff our own warehouses—two facilities in Richmond, CA. We are developing systems to ensure the efficient preservation of millions of books for the long-term, under the leadership of **Director, Kelly Ransom**.

I. Researching, Testing, and Validating Financial Sustainability Models

Even before we start, the Internet Archive and our evaluation partner, New Knowledge, are planning how to sustain and grow the *Open Libraries* assets. Preservation is one of our most critical goals, and sharing the assets across a wide network of libraries is one of our fundamental strategies. As our advisors, LOCKSS founders, Victoria Reich and David Rosenthal would attest, “*Lots of Copies Keep Stuff Safe*.”

The good news is that this project provides for the long-term file maintenance and storage of millions of books. That means we are committed to serving the files in myriad formats in decades to come, just as we do now for billions of web, video, audio, and software files. Planning for the long-term is part of our DNA.

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

In the library world, OCLC is among the most adept at turning cooperative library services into self-sustaining businesses. They've done this for digital card catalogs, interlibrary loans, and now perhaps for e-delivery of the books themselves. We will work with OCLC to (#1) conduct market research and draft business plans for three business opportunities we've identified. Based on the results, we will pilot the most promising models (#2). If the pilots prove to be successful, we will refine our plans and launch a fee-based service (#3) to provide a sustainable source of revenue for the legacy phase of this project. These explorations will be led by our **Books Business Development Manager**.

IA has launched and runs two successful fee-based services: our web archiving business, Archive-It, and our book scanning services offered at 20 locations around the world. (See Financial Strength and Stability section.)

J. Training and Library Support

Our web archiving service, Archive-It, deploys a team of partner specialists—both librarians and engineers—who work with 500 partners for onboarding, training, and remediating glitches in the system. They run an efficient, supportive, community-oriented service, that emphasizes partnership over profit. With *Open Libraries* we will emulate our web team's model of training and support (#1-2). We will create online resources and tutorials to reduce the costs of direct customer service (#3). This effort will be led by our **Partner Support Manager**.

Over the last decade, our Books Team has successfully served more than 1000 contributing libraries who worked with us to scan their collections. We are digitizing the books of 272 active partners across 20 scanning centers today. Expanding to a digitize-and-lend service is a natural evolution of our model of financially sustainable social entrepreneurship.

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

K. Measurement & Evaluation for Accountability, Feedback, and Recalibration

In selecting partners for measurement and evaluation, we searched for several things. We looked for team members who are whip-smart, data-driven, and who know the landscape of libraries well. We wanted our efforts to help not only the Internet Archive, but the entire library ecosystem, which is transitioning through radical technological change. If we were going to devote a substantial part of the budget to these activities, we hoped to leave behind an enduring asset that lives beyond the five-year project span. Thanks to program officers at the Bill and Melinda Gates Foundation, who have run the Global Libraries division for two decades, we were introduced to our two measurement and evaluation teams.

Leading our evaluation team is the New Knowledge Organization, directed by social scientist, Dr. John Fraser, and guided by Beverly Sheppard, former director of the Institute of Museum and Library Services (IMLS). We appreciated that they have been instrumental in creating strategies for the ALA and PLA on measuring the impact of libraries' public programming. Fraser is also the Editor-in-Chief of *Curator: The Museum Journal*, a Tier 1 journal published by Wiley; he has witnessed first hand the journal world's rapid shift to e-delivery. Understanding the intersection between publishers and libraries may be critical to our success. New Knowledge is designing a process that will challenge our direction at key intervals and help us to remain responsive, iterative, and true to our impact goals (#2).

Joining them are two research centers within the University of Washington's (UW) Information School: the Technology and Social Change (TASCHA) group and the DataLab. As the Gates Foundation winds down its Global Libraries program in 2018, it invested in three legacy partners, giving them

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

each a decade of funding to create a “strong and resilient” library field.

TASCHA is one of those three, working to support libraries as they transform into “critical centers of learning, creativity, and community development.”

Joining them are the data scientists of the DataLab, led by Dr. Carole Palmer.

We believe the University of Washington Information School is strategically placed to “direct and influence the practice of librarianship and the role that libraries must play in the lives of people and communities in the 21st century,” as Dean Harry Bruce writes in his letter of commitment to this project.

We’ve asked TASCHA and the DataLab to create the **Impact Data Trust**, a transparent, public repository for library user data to be stored in the Internet Archive (#1). By aggregating use data from IA and all its partners, data scientists can glean more accurate analyses, not only for us, but for the field. The idea for this trust came from IA Advisor, Dr. Susan Hildreth, former director of IMLS and Professor in Practice in UW’s Masters of Library and Information Sciences (MLIS) program. She knows that libraries have long desired a trusted place to aggregate data, in order to see the big picture and make compelling cases to agencies that fund them. But librarians also require the highest ethical standards around reader privacy. Dean Bruce writes about this project hitting a “sweet spot” where data scientists and library practitioners can “leverage data for the social good, in an ethical manner that can inform policy and impact lives for the better.” That is indeed our goal.

TASCHA will work with New Knowledge to convene a yearly meeting of practitioners to share their data-driven insights (#3) and explore field-wide understanding of how *Open Libraries* is shifting use and access. The results of these meetings will inform our decisions in awarding evaluation sub-contracts to field-specific experts to conduct focus groups, surveys, and research that respond directly to questions as they arise. This plan

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

focuses on three goals: accountability to our targets (#1), ongoing data-driven feedback for quick course correction (#2), and producing data-driven lessons that strengthen not only us, but the entire ecosystem we serve (#3).

L. Outreach and Marketing to Key Influencers and Communities

While we work each day with technology, libraries, and books, it is people who are the heart of our project. Among them are half a billion people with disabilities that impact reading, the two-thirds of Americans who have library cards,⁷ the 42 million people who have read a book online at one of our websites, archive.org and OpenLibrary.org in the last six months alone, another 319 million IA patrons who downloaded the text to read on their own devices in 2017, and above all, there are the individuals, who are enriched, educated, and empowered by the trusted information safeguarded by our libraries.

To capture the spirit and purpose of *Open Libraries*, we are launching OpenLibraries.online (#1), a website highlighting a few of these first person stories. We didn't have to go far. Every day our patrons send us their real life experiences. We talked to our mothers, our children, and our colleagues—all of whom will explain to you firsthand why having knowledge at their fingertips is so important. How ebooks have been game changers for them. How the public library is their only technical access point, and how having ebooks there that reflect their own stories makes all the difference between feeling included, or shut out.

Today, a lot of people think that Amazon has made libraries obsolete. But for-profit platforms including Google and Amazon make no guarantees

⁷ American Library Association, Library Cards: Home
<http://libguides.ala.org/librarycards>

**Five-Year
Plan for Open
Libraries—
a project by
the Internet
Archive**

about reader privacy, long-term public access, or preserving the best books produced by humankind. As John Palfrey writes in his book, *BiblioTech: Why Libraries Matter More than Ever in the Age of Google*, “Librarians worry, with some reason, about what protections readers will have when the police come calling for information about the books they have checked out...Their fear is justified: for-profit firms may not stand as firm as libraries would in the face of state pressure.”⁸

So we need to remind Americans why libraries matter and that our world is split by an informational divide. We’ll do that with our first digital media campaign, “Everyone Deserves to Learn,” a primer in the ways ebooks can be lifelines for certain readers (#2). We will follow that with a second campaign, “Inspired to Lead,” about the remarkable leaders who took a stand after internalizing a lesson they discovered in a book. Librarian of Congress, Carla Hayden, often quotes Frederick Douglass who said, “Once you learn to read, you will be forever free.” “If you can absorb information yourself and make your own decisions, that’s a freedom,” Hayden explains.⁹ In the midst of the Baltimore riots protesting Freddie Gray’s killing, Hayden made her own decision: to keep a branch of her library open as a refuge. The power of an idea gleaned from history, translated into action. In the right hands, that is what a book empowers.

Our third campaign will engage directly with the public, encouraging them work with us to make this digital equivalent of a “great public library” a collective endeavor.

⁸ John Palfrey, *BiblioTech: Why Libraries Matter More than Ever in the Age of Google*, (New York: Basic Books, 2015), 204

⁹ Hayden, Carla. “Carla Hayden Thinks Libraries Are a Key to Freedom.” *The New York Times*, January. 19, 2017 <https://www.nytimes.com/2017/01/19/magazine/carla-hayden-thinks-libraries-are-a-key-to-freedom.html>

**Five-Year
Plan for *Open
Libraries*—
a project by
the Internet
Archive**

Finally, our team members are fanning out far and wide to share the *Open Libraries* story (#3). We start this August in front of 2000 Wikipedians in Montreal and move on this Fall to 1800 librarians and publishers in Charleston. Throughout the next five years, we'll be telling our story, listening to our stakeholders, learning from our setbacks, and staying laser-focused on our mission:

When *Open Libraries* is complete—

...access to knowledge will be more evenly distributed, especially to people in marginalized communities around the globe.

...the US library system will be transformed, offering nearly equal numbers of ebooks and hard copies, freeing up space and money so librarians can spend their resources providing critical services to patrons.

...the ebooks on library shelves will look more like the people who read them.

...a century of missing books will be preserved, held in dozens of institutions across borders.

...people in search of information they can trust will find it in the place they've always trusted: their local library.

Open Libraries Project Plan

Summary View

| Task Name | Duration | Start Date | End Date | Assigned To |
|---------------------------------------------------------------------------------------------------------------------------------------------------|----------|------------|----------|------------------------------------|
| FIVE-YEAR PLAN FOR OPEN LIBRARIES A Project of the Internet Archive | | | | |
| A. Governance: Establish teams for guidance, accountability, reporting | 1540d | 10/11/17 | 09/05/23 | Brewster Kahle |
| 1. Establish Advisory Board in seven categories: Tech; Curation; Collection Development; Legal; Service Plan; Accessibility & Financial Modelling | 1313d | 10/11/17 | 10/21/22 | Wendy Hanamura |
| 2. Internet Archive Board Review Scheduled | 1041d | 12/07/18 | 12/02/22 | Jacques Cressaty |
| 3. Internet Archive External Audit scheduled | 1113d | 06/01/19 | 09/05/23 | Jacques Cressaty |
| 4. Internal Grant Financial Management Quarterly and Annual Reviews | 1272d | 04/16/18 | 02/28/23 | Scott Fong |
| 5. MacArthur Grant Report & Internal Assessment (Half-yearly) | 1208d | 05/21/18 | 01/04/23 | Grant Manager |
| B. Hire team to execute Open Libraries | 1221d | 05/01/17 | 01/03/22 | BZ Petroff |
| C. Curation | 1305d | 01/01/18 | 12/30/22 | Curator of Books |
| 1. Establish College Core List | 1220d | 04/30/18 | 12/30/22 | Curator of Books |
| 2. Establish Public Library Core List | 1220d | 04/30/18 | 12/30/22 | Curator of Books |
| 3. Establish Bookshare List for People with Print Disabilities | 1304d | 01/02/18 | 12/30/22 | Curator of Books |
| 4. Establish Most Widely Held by Libraries List | 1240d | 04/02/18 | 12/30/22 | Curator of Books |
| 5. Establish Wikipedia Books List | 1096d | 10/19/18 | 12/30/22 | Curator of Books |
| 6. DPLA Curation Corps selects 1M books for K-16 Education & Diversity | 1302d | 01/01/18 | 12/27/22 | Curator of Books |
| 7. Curate List for People with Disabilities in Marrakesh nations | 783d | 01/01/20 | 12/30/22 | Curator of Books |
| 8. Digital Library Federation (DLF) leads Inclusive Curation project to oversee diverse selection with community feedback | 1304d | 01/01/18 | 12/29/22 | Bethany Nowvskie |
| D. Book Sourcing—develop books pipeline for purchase and digitization | 1175d | 07/02/18 | 12/30/22 | Director of Books |
| 1. Negotiate scanning agreements with partners for Super Scanning Centers (e.g., Univ. of Alberta; Northeastern) | 180d | 01/01/19 | 09/09/19 | Director of Books |
| 2. Negotiate terms of purchase with publishers, vendors, authors, booksellers | 1044d | 01/01/19 | 12/30/22 | Books Business Development Manager |
| 3. Negotiate with University Presses and other publishers to digitize their backlist books | 522d | 07/02/18 | 06/30/20 | Books Business Development Manager |
| 4. Create a pipeline for quality, open access books by global NGOs | 1044d | 01/01/19 | 12/30/22 | Director of Books |

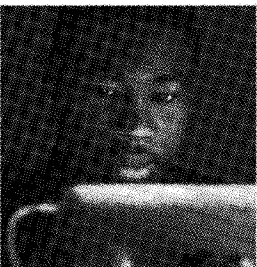
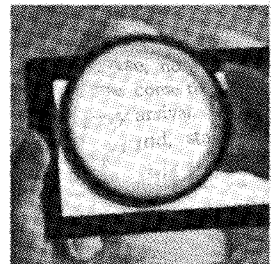
29

| Task Name | Duration | Start Date | End Date | Assigned To |
|--------------------------------------------------------------------------------------------------------------------------------------------|----------|------------|----------|------------------------------------|
| 244 | | | | |
| 245 E. Technology Plan | 1566d | 01/01/17 | 12/30/22 | |
| 246 1. Integrate purchased eBooks (from publishers, authors, etc.) | 1304d | 01/01/18 | 12/29/22 | |
| 247 2. Scale Book Digitization Capacity in Scanning Super Centers | 1566d | 01/01/17 | 12/30/22 | Head of Digitization |
| 254 3. Improve Internet Archive books for use by people with print disabilities(e.g. OCR, chapters, formats) | 720d | 05/01/17 | 01/31/20 | Hank Bromley |
| 255 4. Improve accessibility of web site for patrons with print disabilities | 360d | 06/01/18 | 10/17/19 | Brenton Cheng |
| 256 5. Increase circulation of ebooks to US and Global blind and print disabled communities | 1128d | 09/03/18 | 12/28/22 | Brenton Cheng |
| 260 6. Develop and implement technology for Circulation and Lending (e.g. OPDS, ILS and OPAC integration) | 1179d | 01/01/18 | 07/07/22 | Brenton Cheng |
| 261 7. Integrate capability to cite directly from IA's ebooks with top sources for online information (e.g.--Wikipedia, OER Commons, etc.) | 923d | 06/01/19 | 12/13/22 | Brenton Cheng |
| 262 8. Enable (non-consumptive) Research Use of the Digital Book Dataset | 360d | 06/01/18 | 10/17/19 | John Gonzalez |
| 263 | | | | |
| 264 F. Build Partnerships with Key Stakeholders | 1480d | 05/01/17 | 12/30/22 | Wendy Hanamura |
| 265 1. Build Curation Partnerships | 1298d | 01/02/18 | 12/22/22 | Wendy Hanamura |
| 270 2. Secure Support and Guidance from Organizations Serving People with Disabilities | 1274d | 01/02/18 | 11/18/22 | John Gonzalez |
| 276 3. Secure Support and Guidance from Partnerships for Education & Open Knowledge | 120d | 01/02/18 | 06/18/18 | Mark Graham |
| 280 4. Engage with Technology Providers | 1305d | 01/01/18 | 12/30/22 | John Gonzalez |
| 291 5. Engage Early Adopters & Archetype Innovators | 1302d | 01/02/18 | 12/28/22 | Jim Michalko |
| 301 6. Establish network of Key Influencers in Library Ecosystem | 243d | 05/01/17 | 04/04/18 | Jim Michalko |
| 309 | | | | |
| 310 G. Establish Legal Frameworks and Agreements | 1461d | 05/23/17 | 12/27/22 | Lila Bailey |
| 311 1. Create group of Copyright Scholars to advise on legal framework for D&LS | 1413d | 05/23/17 | 10/20/22 | Lila Bailey |
| 316 2. Legal Working Group continues to expand scholarship, disseminate learnings in wider ecosystem | 1045d | 10/19/18 | 10/21/22 | Lila Bailey |
| 322 3. Create contracts and legal guidelines for compliance with Copyright in D&LS | 1297d | 01/08/18 | 12/27/22 | Lila Bailey |
| 325 | | | | |
| 326 H. Oversee Physical Storage, Inventory and Tracking of assets | 1305d | 01/01/18 | 12/30/22 | Kelly Ransom |
| 327 1. Establish staffing, intake systems, pallet tracking plans to store 3+ million books | 1305d | 01/01/18 | 12/30/22 | Kelly Ransom |
| 333 | | | | |
| 334 I. Financial Sustainability Model | 1030d | 08/30/19 | 08/10/23 | Books Business |
| 335 1. Conduct Market Research and draft Business Plan for earned income from D&LS | 240d | 08/30/19 | 07/30/20 | Books Business Development Manager |

| Task Name | Duration | Start Date | End Date | Assigned To |
|-----------------------------------------------------------------------------------------------------------------|----------|------------|----------|-------------------------|
| 340 2. Pilot most promising option(s) | 240d | 07/31/20 | 07/01/21 | Books Business |
| 343 3. Launch Service of most promising pilot | 200d | 07/02/21 | 04/07/22 | Books Business |
| 344 4. Analyze sustainability model for D&LS based on pilot market data | 90d | 04/08/22 | 08/11/22 | Books Business |
| 345 5. Continue to build fee-based service into sustainable source of revenue | 260d | 08/12/22 | 08/10/23 | Director of Books |
| 346 | | | | |
| 347 J. Training and Partner Support: create systems to on-board, train and support Library Partners | 657d | 07/01/19 | 01/04/22 | Partner Support Manager |
| 348 1. Build Partner Support Team | 657d | 07/01/19 | 01/04/22 | BZ Petroff |
| 353 2. Create plan for partner support, growth & sustainability | 120d | 07/01/19 | 12/13/19 | Partner Support Manager |
| 355 3. Create materials for training and service | 360d | 01/07/20 | 05/24/21 | |
| 359 | | | | |
| 360 K. Measurement & Evaluation for accountability, feedback and recalibration | 1295d | 01/05/18 | 12/22/22 | Director of Books |
| 361 1. Create and run the Impact Data Trust (Trust) | 1285d | 01/06/18 | 12/08/22 | TASCHA |
| 395 2. Metrics, Evaluation and Reporting | 1295d | 01/05/18 | 12/22/22 | New Knowledge |
| 483 3. Community Engagement with Open Libraries' Learnings: Data Mobilization Discussions hosted by UW I-School | 1128d | 05/18/18 | 09/13/22 | TASCHA |
| 506 | | | | |
| 507 L. Outreach and Marketing to Key Influencers & Communities | 807d | 06/01/17 | 07/03/20 | Wendy Hanamura |
| 508 1. Launch Project website: OpenLibraries.online | 97d | 06/01/17 | 10/13/17 | Alexis Rossi |
| 509 2. Digital Media Campaigns for Public engagement | 807d | 06/01/17 | 07/03/20 | Communications Manager |
| 513 3. Public Presentations to Key Influencers & Communities | 101d | 06/24/17 | 11/10/17 | Brewster Kahle |
| 519 | | | | |

30

B. Community Engagement



Stakeholder Identification

At the Internet Archive we are a conduit connecting learners with the published works of humankind. Like the internet itself, we're part of the infrastructure delivering the power of ideas to knowledge seekers and providers. Who benefits from the free flow of ideas? Who has a stake in keeping the power going?

Just about everyone.

Open Libraries will empower **119,000 US libraries**: academic and public library systems, along with the **373,000 library professionals** and associations who serve them.¹ Surveys show when a library offers ebooks, more patrons visit and more books circulate,² but ebooks represent only 17% of their holdings.³ Meanwhile, libraries are bursting at the seams—pressured to preserve ever growing numbers of books, while opening up spaces for public programming. Ebooks enable efficient preservation and community-centric spaces.

Content creators have a huge stake in seeing their works bought and read. Some **authors and publishers** have joined *Open Libraries*, to extend their online reach. The Author's Alliance writes, "Keeping valuable knowledge and creativity out of the reach of readers benefits no one. For authors, the lack of an online version of their books can be tantamount to oblivion."⁴ Other authors and publishers, along with their trade associations, question the idea that libraries' digital lending is fair use. We acknowledge that tension, but believe we can demonstrate that Open Libraries controls circulation in ways greatly benefitting readers without harming creators.

People with disabilities are stakeholders and beneficiaries of our project. They describe a great global "book famine"⁵ and say this effort would open the doors to a world-class accessible digital library.

**Open Libraries**

a project by the Internet Archive

Internet Archive

100&Change

Educators, journalists, scholars and Wikipedia's 116,000 active editors worldwide⁶ all share a stake in preserving public trust in information during this "Post-Truth" era. Our project links information to the trustworthy vetted sources directly in books.

And the beneficiaries?

Every **school, university or public library patron**, including the people who physically visit public libraries 1.5 billion times a year.⁷

More than a **billion people with print disabilities**, including **285 million with visual impairments**.⁸

Online readers whose habits are increasingly monitored and monetized. Librarians staunchly defend reader privacy in stark contrast to for-profit alternatives.

Underserved communities in the US, who will see books reflecting their own lives in libraries.

Wikipedia readers who view **15 billion pages** every month.⁹

College students who could replace costly textbooks with Open Educational Resources, saving on average an estimated \$1000 or more per year.¹⁰

Scientists using computational analysis of books' data to fuel discovery.

Authors whose creative output will be discovered by future generations. As historian, Abby Rumsey, writes, "In 20 years, if a collection cannot be discovered through a web search, people will effectively not know it exists."¹¹

American Democracy which cannot flourish without informed citizens.

Ultimately, we all have a moral stake in connecting the citizenry to freely-accessible knowledge. As Georgetown Law’s Michelle Wu writes, this project “holds the promise of bridging one of the most dangerous divides of our time: an informational divide caused by the growing economic inequality in the United States.”

Inclusiveness

We know that access to knowledge is not evenly distributed. Surveys show that for many marginalized communities, distance, time, cost, and disability pose barriers to knowledge.¹ *Open Libraries* seeks to eliminate those barriers through e-delivery of the critical information locked in books. Working with the National Federation for the Blind (NFB), Benetech, and others, IA will enable people with print disabilities around the world to access a collection of **6.5M ebooks—10x the size of the world’s largest accessible library today.**

A 2015 Pew survey reveals “lower-income Americans, Hispanics, and African Americans are more likely to say that libraries impact their lives and communities than other Americans.” The most frequent library visitors are Hispanics: 21% say they go to the public library at least once a week, versus 14% of all respondents.² Thus, increasing libraries’ e-resources and communal spaces will have outsize impact for these communities.

For readers at the 80% of public libraries that are small and rural,³ we’ll expand access to more ebooks, **from an average of 17% in 2013 to 80% in 2023.** This enables anyone with an internet connection or mobile device to access a wider range of library information, at any time, from anywhere.

Finally, like many institutions, US libraries have helped perpetuate discrimination—sexism, racism, classism—by curating collections based

on values and norms that reflect the biases of the majority population. We'll curate digital collections as diverse as our readers through a transparent, inclusive selection process, with public feedback loops from diverse community panels.

Persons With Disabilities

Less than 10% of published works can be read by people with viewing impairments.¹ To address this book famine, we're teaming with Benetech to produce 10-15,000 fully accessible member-requested books.

Benetech will use its long-tail user data to create lists of books tailored to the needs and desires of people with disabilities. Then we'll post these lists on the Digital Library Federation's dashboard, inviting public feedback. We believe this transparent, community-responsive process can serve as a model for selecting more inclusive collections everywhere.

Our most ambitious plan ensures that everybody in the world with a qualifying disability can access both the IA and Bookshare platforms in a pay-what-you-can model; for most, access will be free. Benetech will launch a Latin American accessible book portal in the 12 Marrakesh Spanish-speaking nations—10 of which have no accessible ebook platforms at all.

We're partnering with organizations worldwide to provide one-click authorization for their certified print disabled members on archive.org, allowing them to download 6.5 million books in formats compatible with adaptive devices. The Accessible Digital Library of India will deliver these ebooks to the world's largest blind population.

Together, the e-collections of these organizations now total 699,000 accessible books—akin to a small college library. We can increase that collection 10x—providing them with the equivalent of a major urban public

library. As the president of the National Federation of the Blind wrote, this project “constitutes the greatest single increase in accessible materials for the blind since... 1931.”

Stakeholder Engagement

At our 2015 Library Leaders Forum,¹ we asked 80 partners this provocative question: *What if IA could wave a wand over your collections and make them digital? Would you lend them in keeping with traditional practices?*

Many agreed it would transform their libraries, unlocking the true value of their collections. But they were concerned they couldn’t explain this conception of fair use to their lawyers.

Library law panelist, Mary Minow, walked us through fair use’s application to controlled digital lending (CDL); Georgetown Law volunteered to launch the first pilot.

University of Alberta’s Gerald Beasley came away inspired, working to clear a path to digitize 500,000 modern books later this year.

In 2016, Georgetown’s Michelle Wu, published her groundbreaking article,² detailing the legal framework for *Open Libraries*. Meanwhile, Kahle travelled the world, sharing our vision with dozens of CTO’s, university leaders, and librarians. Interest was building.

But the costs, the lack of infrastructure, and potential risks made progress slow. We knew creators might question this project, so we began directly engaging with authors and publishers, answering their questions. Authors Alliance members signed on, and we’re digitizing their books for open access. MIT Press Director, Amy Brand, wanted to take control over her titles, after discovering thousands were circulating unauthorized on the web.